ChemPubSoc Europe

DOI: 10.1002/chem.201700957

CHEMISTRY
A European Journal
Full Paper

DNA recognition

Amino-Acid-Derived Naphthalenediimides as Versatile G-Quadruplex Binders

Dora M. Răsădean,[a] Bin Sheng,[a] Jyotirmaee Dash,[b] and G. Dan Pantoş*[a]

Abstract: The design and synthesis of water soluble, amino-acid-functionalised naphthalenediimides (NDIs) as potential ligands of native G-quadruplexes is reported. The NDIs were tested on a panel of oncogene promoters, on the human telomeric sequence *h-telo*, and on double-stranded DNA. Out of the ligands tested, NDI **3** (N_ε-Boc-L-lysine-protected NDI) exhibited a highly discriminating nature by only stabilising the oncogene promoter *c-kit2*, which is up-regulated up to 80% in ovarian, gastrointestinal, and breast malignancies.

Introduction

DNA is a pivotal target in anticancer therapy, and a large number of clinically approved drugs directly interact with it.[1] Of the nucleobase units that comprise the human genome, guanines (G) have the propensity to self-assemble into higher-order structures known as G-quadruplexes.[2] The constitutive element of these stacked G-rich structures is a relatively planar configuration of four guanines stabilised by a cyclic Hoogsteen hydrogen bonding network.[3,4] G-quadruplex DNA displays polymorphic structural forms, which show three major folding topologies that are distinguished based on their strands' orientation: parallel, antiparallel, or a hybrid conformation of the two (Figure 1). The complexity and intricacy of these topologies is highlighted by their ability to form either intra- or intermolecular G-quadruplex motifs with loops and grooves that confer binding preferences.[5]

Recent studies indicate that these secondary structures are encountered in gene bodies.[6] They are also present in other key targets of anticancer therapy, such as introns and double-strand breaks.[1,5] G-quadruplexes have been postulated to govern several crucial biological processes within the chromosome and to regulate transcription, translation, and replication.[1,7–10]

Generally, over-expression of the telomerase enzyme in tumours leads to chaotic and unlimited elongation of the telomere. Thus, anticancer therapies aim to reduce telomerase activity; this leads to stabilised telomeric G-quadruplexes.[1] Such approaches could initiate generate uncapping processes, chromosomal fusions, and senescence, which would eventually trigger DNA alteration and apoptosis.[11,12]

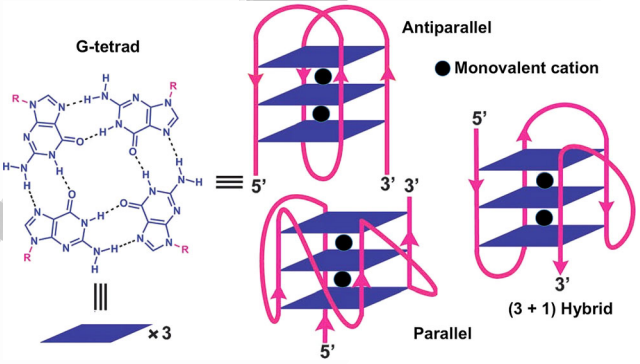


Figure 1. Schematic representation of the structural unit of a G-quadruplex (held together by a cyclic hydrogen-bonding network) and its polymorphic nature.[10]

Owing to the unique topology of each G-quadruplex structure, the prospect of generating highly selective and effective molecules towards the stabilisation of G-rich DNA over dsDNA is promising, but challenging.[13] A relatively large number of small ligands, such as Telomestatin, Quarfloxin, prophyrinoids, amido-anthraquinones, quindoline, and acridine derivatives,[6,14–38] that interact with these higher-order structures have been developed, several of which have entered clinical trials.[1,39,40] Figure 2 illustrates the various G-quadruplex binding modes adopted by small ligands; an example not included in this Figure, which stands out as a unique binding mode, is the threading and stacking interaction, whereby a polyaromatic moiety that bears an alkyl triamine group replaces the alkaline cations and stacks on the top G-tetrad.[29] Naphthalenediimide(NDI)-based compounds have been shown to have a high affinity for G-quadruplexes and have inhibited the telomerase activity at telomeres as a result of *h-telo* stabilization. The success of these ligands is due to their extended planar heteroaro-

[a] D. M. Răsădean, B. Sheng, Dr. G. D. Pantoş
Department of Chemistry
University of Bath
Claverton Down, Bath, BA2 7AY (UK)
E-mail: g.d.pantos@bath.ac.uk

[b] Dr. J. Dash
Department of Organic Chemistry
Indian Association for the Cultivation of Science
2A & @B Raja S C Mullick Road, Kolkata 700032 (India)

Supporting information and the ORCID identification number for the author of this article are available on the WWW under https://doi.org/10.1002/chem.201700957.

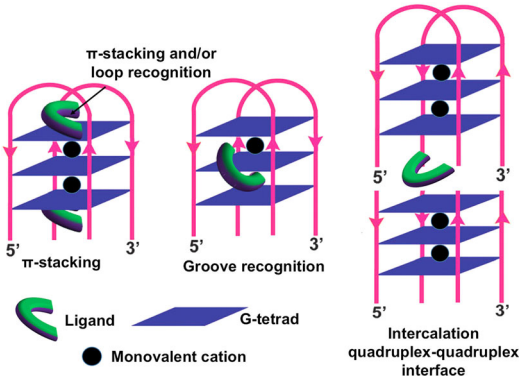


Figure 2. Schematic representation of the principal binding modes to G-quadruplexes adopted by ligands.

matic surface, which facilitates aromatic stacking interactions with the G-tetrads.

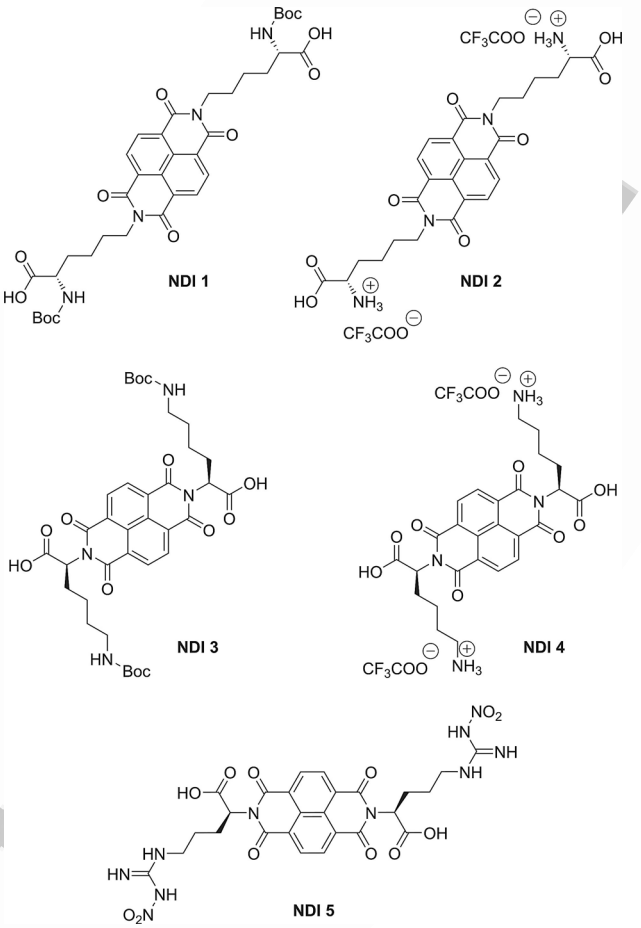
From a synthetic perspective, the fused π-rings of the NDI core provide an excellent platform for the subsequent attachment of scaffolds that can electrostatically interact with the grooves.[41] To date, the NDI-based G-quadruplex ligands include di or tetrasubstituted derivatives that bear either aromatic or aliphatic side chains.[41,42]

Herein, we report the stabilisation potential of certain amino acid functionalised NDIs against a panel of G-quadruplex sequences, which include the human telomeric *h-telo* and the oncogene promoters *c-Myc*, *c-kit*, and *k-ras* as well as a double-stranded DNA (dsDNA). To investigate the thermodynamic driving forces that dominate NDI–G-quadruplex DNA stabilisation and their susceptibility to self-assemble into supramolecular architectures, we conducted temperature-dependent CD spectroscopy.

Results and Discussion

Ligand and condition optimisation

The symmetrical, disubstituted NDI derivatives (**1**–**5**, Scheme 1) were designed for optimum pharmacological behaviour. The planar NDI core can stack on the surface of a guanine tetrad through aromatic stacking interactions, and the amino acid side chains were introduced to increase the water/buffer solubility and interaction with the DNA loops. Lysine and arginine were selected as the amino acid side chains owing to their structure and complementary charge to the G-quadruplex backbone.[1,41–44] The length of the side chain plays a key role in tethering the G-quadruplex DNA binding ligands; a good correlation between shorter linker chains and their affinity for G-quadruplex DNAs over duplex DNA has been reported.[44] Additionally, the attachment of cationic charges on the peripheral substituents increase the aqueous solubility and is likely to influence the binding selectivity. The ligands have been designed as pairs of the same amino acid (N_α-L-lysine and N_ε-L-lysine): the Boc-protected form and its ammonium trifluoroacetate counterpart for NDIs **1**–**4**, whereas NDI **5** is the sole arginine functionalised compound of the series.[45]



Scheme 1. Structures of the NDI derivatives employed in this study.

The synthesis of the symmetric amino acid functionalised NDIs has involved a microwave-assisted strategy, which was based on a protocol that we have previously reported.[46,47] We employed naphthalenetetracarboxylic dianhydride and the corresponding protected amino acids as starting materials for NDIs **1**, **3**, and **5**. NDIs **2** and **4** were obtained by deprotection[48] of NDIs **1** and **3**, respectively (Scheme 1).

These ligands are water soluble, as indicated by linear Lambert–Beer plots of the absorption of solutions. An exception is NDI **2**, which has been shown to aggregate in aqueous solutions at concentrations above 1.9 mM.[49] In light of this, all of the DNA binding studies assumed a non-aggregated state for all ligands except NDI **2**. In this case, the T_m values reflect not only the (de)stabilisation of the quadruplex sequence but also account for the de-aggregation of NDI **2** (a process inherently unfavourable under these conditions).

Fluoride salts, unlike chloride or bromide ones, are transparent in the far-UV and VUV■■please define■■ region; therefore, potassium fluoride was selected as a cation source, which allowed us to monitor the behaviour of the system to wavelengths as low as 184 nm.[50]

CD/ UV/Vis-based structural analysis of the adopted topologies of G-quadruplexes

The CD spectrum is not a reflection of the G-quadruplex DNA base achiral sequence, but rather it is a consequence of the chiral environment induced by the sugar backbone. Therefore, the CD spectrum shows the local structural features of a conformation rather than of the overall sequence.[50,51] The CD fingerprint of each G-quadruplex depends on the glycosyl torsion angles, namely *syn* or *anti*. The latter corresponds to parallel topologies that contain chain-reversal loops; by contrast, the antiparallel-stranded G-quadruplex is characterised by a mixture of both conformers and diagonal loops.[35,52] In general, the CD spectra of parallel motifs display a positive peak around 260 nm and a negative peak near 240 nm, whereas a positive peak at 290 nm and a negative peak at 260 nm are commonly assigned to antiparallel conformations.[52] Figure 3 illustrates the polymorphic nature of G-quadruplex-forming sequences investigated in this work.

All the investigated sequences adopted a parallel topology in PBS (phosphate buffered saline) that contained potassium fluoride (100 mM), except for the telomeric sequence *h-telo*, which folds in an antiparallel manner (Figure 3). Each spectrum of the ligand–G-quadruplex complexes were compared to the CD spectrum of the sequence itself, and as expected, we observed a drop in the relative intensities.

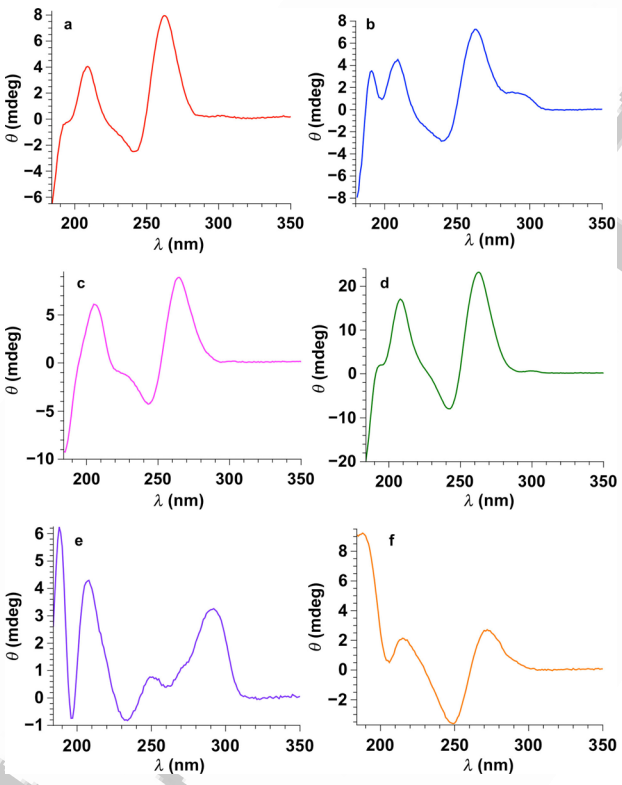


Figure 3. The CD spectra of G-quadruplex sequences involved in this study. Each individual sequence folds in distinct topologies in the presence of K^+ (data collected at 23 °C in 10 mM potassium phosphate buffer and 100 mM potassium fluoride). Unique CD patterns for the parallel conformers, a) 10 μM *c-kit2*, b) 11 μM *c-kit1*, c) 12 μM *c-Myc*, d) 20 μM *k-ras*, as well as the antiparallel assembly e) 9 μM *h-telo*. (f) CD fingerprint of 12 μM dsDNA.

c-kit1 and *c-kit2* are putative G-quadruplex-forming sequences within the same promoter; *c-kit1* is located between −87 and −109 nucleotides, whereas *c-kit2* appears between −140 and −160 base pairs.[13,52] The CD and UV spectra of the *c-kit1* sequence in the presence of NDIs **1**, **3**, **4**, and **5** showed the typical spectral signatures of a parallel-stranded conformation that bears short loops. This is based on the assumption that the loops are too short to interact with opposite guanine units within the same quadruplex,[52] but they can bridge guanines of distinct G-rich strands, which leads to double-chain reversal loops. The CD spectra also revealed a shoulder peak at 295 nm (assigned as an extended tail of the main peak) and two smaller, but well-defined, positive peaks in the far-UV region.

The CD profiles of NDIs **1**, **3**, **4**, and **5** interacting with the *c-kit2* sequence also display a large positive peak at 262 nm (without the shoulder peak), and a less intense negative peak at 241 nm, which indicate a parallel topology with small propeller-like loops. In contrast to *c-kit1*, this sequence exhibits a sole positive peak centred at 207 nm with a poorly defined shoulder around 190 nm. This is slightly shifted upon the addition of ligands; *c-kit1*--NDI **4** assembly shows the most prominent shoulder extension.

The spectral characteristics of molecular assemblies made up of oncogene promoters *k-ras* and *c-Myc* interacting with NDIs **1**, **3**, **4**, and **5** in the presence of K^+ were similar to those highlighted for *c-kit2*; therefore, the peaks were ascribed to a parallel-stranded conformation with short loops. For the *k-ras* sequence, the shoulder peak at 194 nm followed the same trend, with NDI **4** giving the highest positive peak. Distinctively, the CD spectra for the unbound and bound *c-Myc* sequence exhibit a stronger peak around 205 nm with shoulders at higher wavelengths.

Unlike the promoter G-quadruplex-forming sequences, the human telomere *h-telo* incorporates single-stranded overhangs folded in a unimolecular manner.[53] This sequence adopts an antiparallel topology in the presence of K^+, as indicated by a broad positive peak centred at 291 nm and a positive peak near 205 nm. The CD spectrum reveals a sharp positive peak having large ellipticity around 189 nm, which raises the question whether the spectral motif from the far-UV region could serve as a signature for topology identification. The *h-telo* conformation is highly polymorphic, depending on the conditions employed, and is associated with three connecting loops.[54]

As discussed, the appropriate assignment of the pattern identified under 200 nm (far-UV) was carefully conducted to understand whether this is correlated to the conformational fluctuations induced by the binding effect or a fingerprint for the G-quadruplex DNA itself. The CD spectra of the native G-quadruplexes did not vary considerably compared with the spectra of NDI–G-quadruplex assemblies (albeit a small shift of 1–2 nm, particularly for *c-kit1* sequence) and are in good agreement with the CD spectra reported in the literature for the same sequences.[53] This indicates that the ligand has only a small influence on the far-UV peaks that are associated with the quadruplex structures.[53,55]

CD thermal melting studies

The ability of each NDI molecule to selectively stabilise the DNA sequences was assessed by performing variable-temperature CD/UV studies. The melting temperature (T_m) is defined as the equilibrium condition between the folded and unfolded states of a particular assembly.[56] T_m is influenced by multiple factors (i.e., cooperativity of binding, ligand's affinity for a particular conformation, and number of available binding sites), which determine the overall stability of ligand–G-quadruplex interaction.[50] A CD melting spectrum also provides details on the melting profile, as it may be either single-step denaturation or a concerted process that involves several intermediate states.[54]

Thermal stabilisation studies were performed over a broad wavelength range (184–450 nm) in the presence of K^+, and T_m was monitored for each relevant wavelength. The measurements were conducted in mixtures that contained 3 or 10 μM of preannealed DNA and 10 equivalents of NDI.

The T_m value was determined from the temperature versus degree of ellipticity plot by means of singular value decomposition (SVD) and Boltzmann mathematical models; both of which have provided similar results for all sequences. The T_m and ΔT_m values for G-quadruplexes and dsDNA used in this study are summarised in Tables 1–5, which show the temperatures obtained from fitting the Boltzmann model to the experimental data. The analysis of the negative peak near 240 nm displayed large errors for several sequences (*c-kit1*, *c-kit2*, *c-Myc*) owing to a low ellipticity in this region, and the CD signal near 185 nm gave large errors in certain cases, because it approached the lowest detection limit of our CD setup. Therefore, the peaks that corresponded to these wavelengths were excluded from further investigation.

Table 1. T_m and ΔT_m values determined from variable-temperature (VT■■ok?■■) CD studies monitored at a relevant wavelength for the *c-kit2* sequence upon addition of NDIs **1**–**5**. The Boltzman equation was fitted to the experimental data.

Sequence/Assembly	T_m [°C]	ΔT_m [°C]	λ [nm]
c-kit2	78.2 ± 2.7	–	262
NDI 1	88.4 ± 8.4	10.2	264
NDI 2	88.7 ± 7.1	10.5	264
NDI 3	**92.8 ± 5.9**	**14.6**	264
NDI 4	75.1 ± 0.9	−3.1	262
NDI 5	72.3 ± 1.0	−5.9	262

Although the *c-kit2* sequence displayed the highest melting temperature within the G-quadruplexes panel, three out of the five amino acid functionalised NDIs stabilised its structure. NDI **3** showed a strong stabilisation towards this sequence, which was highlighted by a remarkable increase in T_m of 14.6 °C for the peak attributed to G-quadruplex motif at −264 nm (the CD melting associated spectrum for *c-kit2*–NDI **3** is shown in Figure 4; the other associated CD meting spectra are provided in the Supporting Information, Figure S1). NDI **1** and **2** displayed moderate stabilisation, whereas the other ligands exhibited only modest stabilisation or even destabilisation towards this sequence (Table 1).

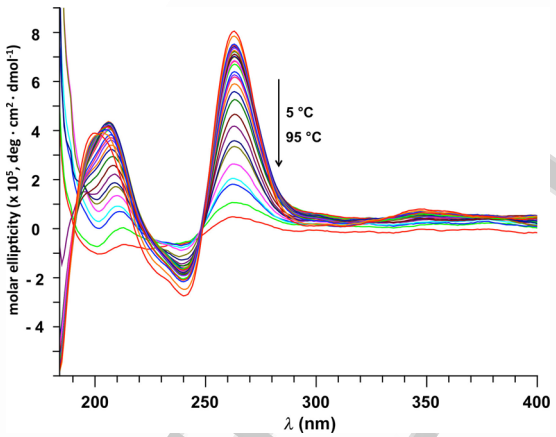


Figure 4. The CD spectra associated with the melting of 3 μM *c-kit2* G-quadruplex sequence that was interacting with NDI **3** (data collected at 23 °C in 10 mM potassium phosphate buffer and 100 mM potassium fluoride).

NDIs **1** and **2** displayed a similar ability to stabilise *c-kit2*; therefore, the presence of the Boc group in the protected NDI **1** did not seem to make a difference in terms of quadruplex stabilisation. However, as will be discussed later, the high T_m that was observed for NDI **2** is not representative of its "real" stabilisation ability owing to its ability to aggregate in solution. Contrary to this, NDIs **3** and **4** displayed an opposite behaviour: the Boc-protected lysine ligand (NDI **3**) provided the largest variation, and thus, the best stabilisation, whereas NDI **4** destabilised *c-kit2*. We postulate that the stabilisation potential of NDI **3** and its selective nature is facilitated by the N_ε-Boc-protecting group, which provides a highly favourable interaction within the quadruplex cavity loops.

The CD thermal studies of *c-kit2*–NDI assemblies revealed an inverted-sigmoid melting profile with isosbestic points around 250 and 220 nm, which indicates a monophasic transition across the studied temperature range. This feature is lacking below the 200 nm region, which shows rather hypsochromic-shifted (blue-shifted) peaks as the temperature increased. Because this region cannot be correlated with the observed topology, such behaviour does not provide valuable information about the process of assembly unfolding during the analysis.

Although G-quadruplex-forming *c-kit1* and *c-kit2* sequences emerged from the same core promoter,[13] the NDI ligands can discriminate between them; *c-kit1* is only moderately stabilised or even destabilised in the presence of our ligands (Table 2). The CD melting associated spectra of *c-kit1* with NDIs **1**–**5** are shown in Figure S2 (see the Supporting Information).

The promoter *c-Myc* was also considerably stabilised by NDI **1** and **2**, as shown in Table 3. The selective stabilisation of this G-quadruplex sequence by NDIs **1** and **2** and not by NDIs **3**–**5** indicates that the steric hindrance/electrostatic requirements represented by the α-COOH group on the side chain adversely influence the interaction. The interaction between NDIs **1** and **2** is most likely through the flat aromatic NDI core, with minor contributions from the aliphatic side chains.

Table 2. T_m and ΔT_m values determined from VT CD studies monitored at a relevant wavelength for the *c-kit1* sequence upon addition of NDIs **1**–**5**. The Boltzman equation was fitted to the experimental data.

Sequence/Assembly	T_m [°C]	ΔT_m [°C]	λ [nm]
c-kit1	65.8 ± 0.7	–	262
NDI 1	67.5 ± 1.0	1.7	263
NDI 2	**68.0 ± 0.8**	**3.0**	262
NDI 3	64.3 ± 0.4	−1.5	263
NDI 4	65.6 ± 0.7	−0.2	264
NDI 5	66.3 ± 0.5	0.5	262

Table 3. T_m and ΔT_m values determined from VT CD studies monitored at a relevant wavelength for the *c-Myc* sequence upon addition of NDIs **1**–**5**. The Boltzman equation was fitted to the experimental data.

Sequence/Assembly	T_m [°C]	ΔT_m [°C]	λ [nm]
c-Myc	63.0 ± 1.5	–	264
NDI 1	72.2 ± 1.5	9.2	264
NDI 2	**74.2 ± 3.8**	**11.1**	265
NDI 3	63.1 ± 2.7	0.1	265
NDI 4	65.2 ± 1.9	2.2	264
NDI 5	63.2 ± 4.1	0.2	264

The propeller-like parallel conformation is retained; the ligands induced only a small shift in the peak pattern, as highlighted for NDI **1** in Figure 5 (see the Supporting Information, Figure S3, NDIs **2**–**5**).

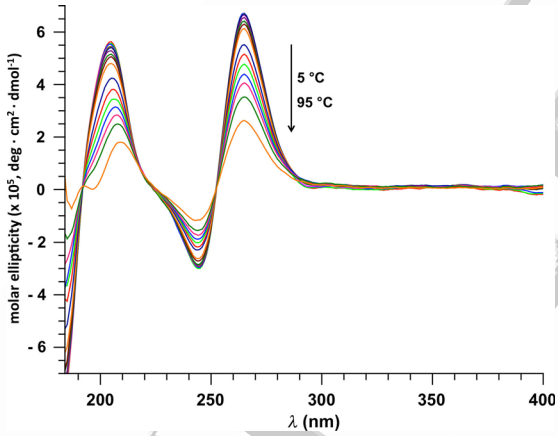


Figure 5. The CD spectra associated with the melting of 12 μM *c-Myc* G-quadruplex sequence that was interacting with NDI **1** (data collected at 23 °C in 10 mM potassium phosphate buffer and 100 mM potassium fluoride).

In terms of oncogene promoter *k-ras* and human telomere *h-telo* stabilisation, NDI **4** displayed the best, albeit minor, stabilisation (4 and 2 °C, respectively; Table 4).

As in the case of *c-Myc*, the parallel and antiparallel topologies did not change significantly as the temperature was increased, but isosbesticity was lost in the far-UV region. The CD melting profile spectra of *k-ras* and *h-telo* interacting with NDI **4** are shown in Figure 6, and the rest of the spectra are given in Figures S4 and S5 (see the Supporting Information).

Table 4. T_m and ΔT_m values determined from VT CD studies monitored at a relevant wavelength for the *k-ras* and *h-telo* (*italics*) sequences upon addition of NDIs **1**–**5**. The Boltzman equation was fitted to the experimental data.

Sequence/Assembly	T_m [°C]	ΔT_m [°C]	λ [nm]
k-ras	58.3 ± 0.6	–	262
h-telo	*62.5 ± 0.3*		*292*
NDI 1	60.6 ± 0.3	2.3	263
	61.8 ± 0.4	*−0.7*	*290*
NDI 2	58.7 ± 1.1	0.4	263
	61.4 ± 1.0	*−1.1*	*290*
NDI 3	60.0 ± 0.5	1.7	263
	61.5 ± 0.6	*−1.0*	*291*
NDI 4	**62.4 ± 0.4**	**4.1**	263
	64.5 ± 0.4	***2.0***	*290*
NDI 5	60.1 ± 0.5	1.8	263
	62.8 ± 0.6	*0.3*	*292*

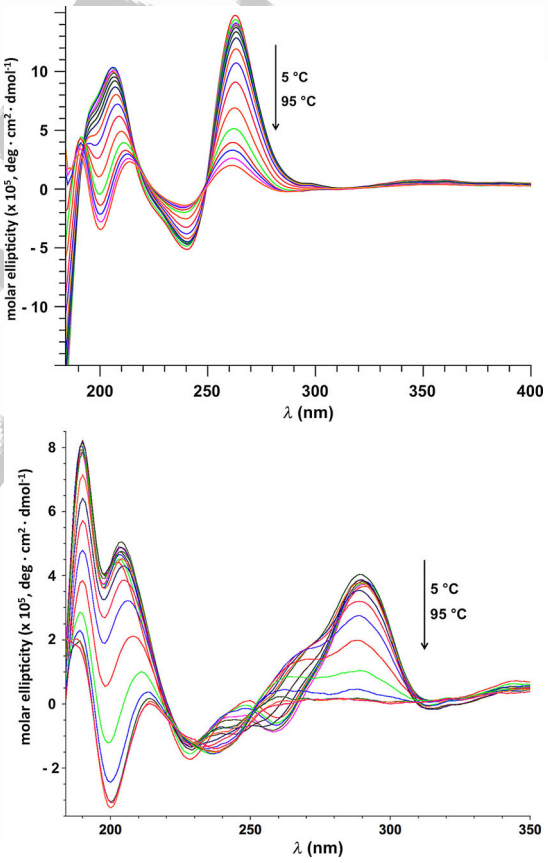


Figure 6. The CD spectra associated with the melting of 12 μM *k-ras* G-quadruplex sequence that was interacting with NDI **4** (top) and 9 μM *h-telo* sequence with NDI **4** (bottom) (data collected at 23 °C in 10 mM potassium phosphate buffer and 100 mM potassium fluoride).

Lastly, the stabilisation effect of NDI ligands on dsDNA was also investigated. The CD melting profile of dsDNA interacting with NDIs **1**–**5** revealed two positive peaks near 270 and 218 nm as well as a negative peak around 250 nm. The far-UV region was dominated by an intense, broad positive peak centred at 193 nm. Similar to the melting CD spectra of G-quadruplexes, the clear isosbestic points observed across the UV

region were missing below 200 nm. Moreover, the CD fingerprint of dsDNA did not change considerably upon addition of the ligand (see the Supporting Information, Figure S6).

The melting studies showed either destabilisation or minor stabilisation for NDIs **1** and **4** (under 2 °C), which is reflected in the T_m values (Table 5). This suggests that our ligands do not undergo a significant interaction with dsDNA, which confirms that these ligands are selective binders of certain G-quadruplex sequences.

Table 5. T_m and ΔT_m values determined from VT CD studies of sequences, which were monitored at a relevant wavelength for dsDNA, upon addition of NDIs **1**–**5**. The Boltzman equation was fit to the experimental data.

Sequence/Assembly	T_m [°C]	ΔT_m [°C]	λ [nm]
dsDNA	55.1 ± 0.8	–	273
NDI 1	**57.0 ± 1.2**	**1.9**	273
NDI 2	55.8 ± 0.7	0.7	273
NDI 3	52.6 ± 0.8	−2.5	273
NDI 4	56.9 ± 1.5	1.8	273
NDI 5	**52.4 ± 1.8**	−2.7	274

An essential question that arises when studying such assemblies is related to the mode of ligand interaction with the G-quadruplex core. With respect to our rational design of NDI derivatives, we assume that non-covalent forces, namely aromatic stacking interactions within the quasi-planar guanine tetrads and hydrogen bonding networks, are the principal binding interactions adopted by these molecules.

The heat map in Figure 7 concisely describes the stabilisation potential of individual NDIs that is induced on each DNA sequence. NDI **3** exhibited the greatest stabilisation towards the *c-kit2* sequence, whereas the same molecule had only a modest effect on the rest of the panel. NDI **5** displayed the least favourable binding potential towards these higher-order structures.

	c-kit2	*c-kit1*	*c-Myc*	*K-ras*	*h-telo*	*dsDNA*
NDI 1	10.2	1.7	9.2	2.3	-0.7	1.9
NDI 2	7	3	11.1	0.4	-1.1	0.7
NDI 3	14.6	-1.5	0.1	1.7	-1.1	-2.5
NDI 4	-3.1	-0.2	2.2	4.1	2	1.8
NDI 5	-5.9	0.5	0.2	1.8	0.3	-2.7

Figure 7. The heat map table, which displays the ΔT_m values at 260, 270, and 290 nm that correspond to each NDI and G-quadruplex assembly investigated in this study. The trend decreases from green to red coding.

The effective impact of NDIs **3** and **1** towards *c-kit2* stabilisation is particularly noteworthy, because this oncogene promoter is overexpressed up to 80% in ovarian, gastrointestinal, and breast malignancies.[13] The roles of the *c-kit* sequence cover essential biological processes, from the regulation of apoptosis and differentiation to cell survival.[13]

Unravelling the behaviour of the interaction between NDI 2 and G-quadruplex DNA

The behaviour exhibited by G-quadruplex–NDI **2** assemblies were studied further owing to the peculiar binding data that was observed for this ligand when compared with the other NDIs employed. The CD spectra of these complexes, regardless of the type of sequence (Figure 8, *c-kit2* interacting with this

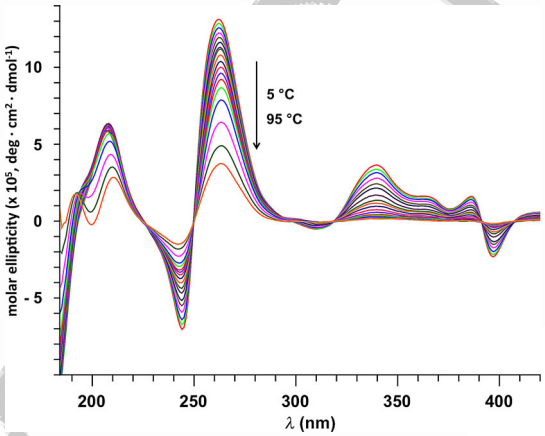


Figure 8. The CD melting spectra of the assembly that is formed between NDI **2** and 12 μM *c-kit2* (data collected at 23 °C in 10 mM potassium phosphate buffer and 100 mM potassium fluoride).

NDI **2**; see the Supporting Information, Figure S7, other G-quadruplex sequences), displayed a CD signal centred at 381 nm, which was ascribed to the NDI core. This led us to propose that this ligand is involved in a binding competition between the self-aggregation process[49,57–62] and the interaction with the G-quadruplex sequences.

To investigate the self-aggregation process, we performed variable-temperature studies on a sample of NDI **2** in potassium phosphate buffer (pH 7.4), as for the DNA–ligand experiments. The CD fingerprint of NDI **2**, the melting curve profile, and thermodynamic functions associated with the process are shown in Figure 9. As determined from fitting the isodesmic polymerisation model, the percentage of monomeric NDI **2** in the solution dominated at low temperature, but this decreased as the temperature increased. At temperatures above 338 K (65 °C), which is the T_m for this molecule, we only observed 50% of monomeric NDI, and the degree of polymerisation was gradually increased.

The positive value of ΔS indicates that an entropically favoured process governs the self-assembly of this molecule. This behaviour is typical for an aromatic stacking/hydrophobic interaction.

Conclusions

The stabilisation of G-quadruplex DNA as the target in anticancer therapy is of broad and current interest, and the design of ligands with unique affinity towards one G-quadruplex-forming sequence is challenging. Out of the five ligands that

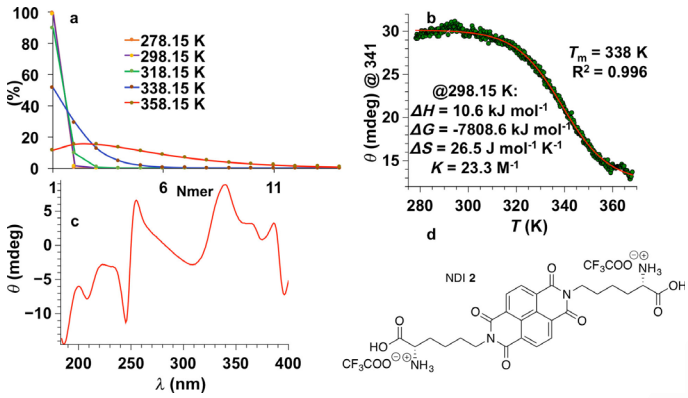


Figure 9. a) The VT CD data of NDI **2**, which fits to the isodesmic model as a function of the degree of polymerisation. b) The Boltzmann fitting of the NDI **2** melting curve (3×10^{-4} M in 10 mM potassium phosphate buffer and 100 mM potassium fluoride) of the maxima at 341 nm as a function of temperature and associated thermodynamic parameters calculated at 298.15 K. c) The CD spectrum of NDI **2** (3×10^{-4} M) at ambient temperature. d) The structure of NDI **2**.

we screened against five biologically relevant G-quadruplexes, NDI **3** only stabilised *c-kit2*, whilst the parallel topology of the sequence was maintained. This discrimination is remarkable, especially because NDI **3** has a small destabilisation effect on *c-kit1*. Somewhat surprisingly, the Boc-protected NDI derivatives stabilise the G-quadruplexes equally well or stronger when compared with their positively charged analogues. This highlights the delicate balance between polarity and hydrophobicity that needs to be achieved for efficient quadruplex binding. Overall, the NDI ligands are superior for stabilising the *c-kit2* and *c-Myc* sequences over *c-kit1, k-ras, h-telo*, and dsDNA. This work paves the way towards a deeper understanding of the structure–activity relationship between NDI ligands and G-quadruplexes. Further work on NDI–G-quadruplex binding studies are currently underway in our group, and it will be published in due course.

Experimental Section

Materials

The guanine-rich oligonucleotides were purchased from Invitrogen®, and the nucleotides order in the sequence were as follows (from 5′ to 3′): *c-kit2*: GGG CGG GCG CGA GGG AGG GG; *c-kit1*: GGG AGG GCG CTG GGA GGA GGG; *c-Myc*: TGA GGG TGG GTA GGG TGG GTA A; *k-ras*: AGG GCG GTG TGG GAA GAG GGA AGA GGG GGA GG; *h-telo*: AGG GTT AGG GTT AGG GTT AGG GT; dsDNA: TAT AGC TAT AHE GTA TAG CTA. The oligonucleotide solutions (3 μM for *c-kit2* and 10 μM for the rest of the sequences) were previously annealed in 10 mM KH_2PO_4/K_2HPO_4 buffer (pH 7.4) and 100 mM KF, and they were used without further purification. The annealed DNA was stored at 4 °C prior to use.

All the other reagents and solvents were supplied by either Sigma–Aldrich, VWR, or TCI. Samples for NMR spectroscopy were made up with [D_6]DMSO. NMR spectra were acquired on a Bruker spectrometer at 300 MHz for ^{1}H NMR and 75 MHz for ^{13}C NMR; they were referenced to the residual solvent peaks. All spectra were recorded at 298 K, and data were reported in ppm. The NSI spectra (negative ion) were recorded on a LTQ Orbitrap XL hybrid FTMS instrument. Microwave-assisted reactions were conducted in a CEM microwave reactor.

Synthesis of symmetrical amino acid functionalised NDIs 1, 3, 5

1,4,5,8-Naphthalenetetracarboxylic dianhydride (NDA) and the corresponding amino acid were suspended in dry DMF (5 mL) in a microwave tube. Dry Et_3N (0.2 mL) was added, and the mixture was sonicated until it became a fine suspension. The reaction mixture was heated under microwave irradiation for either 5 min (NDIs **1** and **3**) or 10 min (NDI **5**) at 140 °C. The solvent was removed under reduced pressure to yield a brown residue, which was further suspended in acetone. The suspension was added dropwise to stirring $HCl_{(aq.)}$ (0.25 M). The resulting precipitate was filtered and dried in vacuum to yield a coloured solid.

NDI 1:[49] The reaction was performed on 1 equivalent of NDA (200 mg, 0.746 mmol) and 2 equivalent of N_α-Boc-L-lysine (367 mg, 1.491 mmol) by using the general procedure. Yield of NDI **1**: 461 mg, 85%; ^{1}H NMR: δ = 12.42 (br ■■s or m?■■indicate range if m, also below■■, 2H,), 8.63 (s, 4H), 7.06 (d, $J_{(H,H)}$ = 7.9 Hz, 2H), 4.03 (4H, t, $J_{(H,H)}$ = 7.2 Hz), 3.86–3.79 (m, 2H), 1.47–■■ ppm (m, 30H); ^{13}C NMR: δ = 174.6, 163.0, 156.0, 130.8, 126.6, 126.5, 78.2, 53.7, 30.84, 28.5, 27.38.0, 23.6 ppm; FTMS-NSI: *m/z* calcd for $C_{36}H_{44}N_4O_{12}$: 723.2856 [*M*−H]$^-$; found: 723.2883.

NDI 3:[46] The reaction was performed on 1 equivalent of NDA (200 mg, 0.746 mmol) and 2 equivalent of N_ε-Boc-L-lysine (367 mg, 1.491 mmol) by using the general procedure. Yield of NDI **3**: 383 mg, 72%; ^{1}H NMR: δ = 12.93 (s, 2H), 8.73 (s, 4H), 6.70 (t, $J_{(H,H)}$ = 5.7 Hz, 2H), 5.50 (dd, $J_{(H,H)}$ = 9.4, 5.0 Hz, 2H), 2.85–2.74 (m, 4H), 2.28–2.12 (m, 2H), 2.12–1.95 (m, 2H) 1.45–1.05 ppm (m, 26H); ^{13}C NMR: δ = 171.0, 162.7, 155.8, 131.6, 126.7, 126.3, 77.5, 53.9, 36.2, 31.1, 29.7, 28.4, 23.6 ppm; FTMS-NSI: *m/z* calcd for $C_{36}H_{44}N_4O_{12}$: 723.2867 [*M*−H]; found: 723.2883.

NDI 5: The reaction was performed on 1 equivalent of NDA (100 mg, 0.373 mmol) and 2 equivalents of H–Arg–(NO_2)–OH (163.5 mg, 0.746 mmol) by using the general procedure. Yield of NDI **5**: 100 mg, 44%; ^{1}H NMR: δ = 12.95 (s, 2H), 8.65 (s, 4H), 8.42 (br■■s or m?■■indicate range if m, also below■■, 2H), 7.76 (br, 4H), 5.47 (dd, $J_{(H,H)}$ = =9.6, 4.7 Hz, 2H), 3.13–2.99 (m, 4H), 2.31–2.07 (m, 4H), 2.07–1.79 (m, 2H), 1.46 ppm (br■■, 4H); ^{13}C NMR: δ = 170.9, 162.8, 159.5, 131.6, 126.7, 126.4, 53.6, 34.5, 26.0, 25.6 ppm; FTMS-NSI: *m/z* calcd for $C_{28}H_{28}N_8O_{14}$: 669.1659 [$C_{26}H_{26}N_{10}O_{12}$−H]$^-$, found: 669.6663.

Synthesis of symmetrical deprotected derivatives NDIs 2 and 4

NDIs **1** and **3** served for the synthesis of the corresponding deprotected NDIs **2** and **4** by using the following general procedure: The starting material was suspended in CH_2Cl_2 (5 mL) and trifluoroacetic acid (5 mL) was added to the suspension. The reaction mixture was stirred at room temperature for 3 h. The solvents were removed under reduced pressure, and the residue was subsequently washed with diethyl ether (2×30 mL), filtered, and vacuum dried to yield a brown powder in both cases.

NDI 2:[49] The aforementioned general procedure was followed by using NDI **1** (350 mg). Yield of NDI **2**: 332 mg, 91%; ^{1}H NMR: δ = 8.62 (s, 4H), 8.27 (br■■, 6H), 4.02 (t, $J_{(H,H)}$ = 7.3 Hz, 4H), 3.96–3.84 (m, 2H), 1.91–1.60 (m, 8H), 1.57–1.29 ppm (m, 4H); ^{13}C NMR: δ = 171.4, 162.9, 158.8, 158.3, 130.8, 126.6, 126.4, 52.1, 50.1, 30.1, 28.9, 27.3, 24.0, 22.3 ppm; FTMS-NSI: *m/z* calcd for $C_{30}H_{30}F_6N_4O_{12}$: 733.1566 [*M*−H_2O−H]$^-$; found: 733.1586.

NDI 4: The aforementioned general procedure was followed by using NDI **3** (350 mg). Yield of NDI **4**: 102 mg, 33%; ^{1}H NMR: $\delta =$ 12.99 (s, 2H), 8.74 (s, 4H), 7.66 (s, 6H), 5.53 (dd, $J_{(H,H)} = 9.0$, 5.2 Hz, 2H), 2.87 (s, 1H), 2.80–2.64 (m, 4H), 2.34–2.17 (m, 2H), 2.13–1.92 (m, 2H), 1.63–1.21 ppm (m, 8H); ^{13}C NMR: $\delta = 171.0$, 162.8, 131.6, 126.7, 126.4, 53.6, 38.9, 28.4, 27.2, 23.2 ppm; FTMS-NSI: *m/z* calcd for $C_{30}H_{30}F_6N_4O_{12}$: 751.1667 $[M-H]^-$; found: 751.1692.

Variable-temperature studies

CD and UV/Vis experiments were performed on an Applied Photophysics Chirascan Circular Dichroism Spectrophotometer equipped with a Peltier temperature controller under the following parameters: wavelength scanning range 184–450 nm; scanning increments 1 nm; monochromator bandwidth 2.5 nm; sampling time-per-point 2 s. The temperature range used for variable-temperature studies was either 5–95 °C with 3 °C increments *(c-kit2)* or 5–90 °C with 5 °C increments (other DNA sequences); temperature equilibration time 45 s. A 1 mm path-length quartz cuvette was used for all experiments. The background spectrum of the buffer was recorded at 23 °C and was subtracted from all measurements. Each measurement required 200 μL of the DNA sequence solutions and 10 equivalents of the ligand. The raw data were processed and analysed by using two different programmes: Qtiplot® and CD Apps.

Acknowledgements

We thank the British Council-DST UKIERI and the University of Bath for financial support and the EPSRC UK National Mass Spectrometry Facility at Swansea University for high resolution analyses. We thank Mr Rakesh Paul for carefully reading the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Keywords: circular dichroism · DNA · g-quadruplex · naphthalenediimide · structure–activity relationships

[1] C. Marchetti, A. Minarini, V. Tumiatti, F. Moraca, L. Parrotta, S. Alcaro, R. Rigo, C. Sissi, M. Gunaratnam, S. A. Ohnmacht, *Bioorg. Med. Chem.* **2015**, *23*, 3819–3830.
[2] D. Rhodes, H. J. Lipps, *Nucleic Acids Res.* **2015**, *43*, 8627–8637.
[3] S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian, S. Neidle, *J. Am. Chem. Soc.* **2005**, *127*, 10584–10589.
[4] G. W. Collie, R. Promontorio, S. M. Hampel, M. Micco, S. Neidle, G. N. Parkinson, *J. Am. Chem. Soc.* **2012**, *134*, 2723–2731.
[5] M. L. Bochman, K. Paeschke, V. A. Zakian, *Nat. Rev. Genet.* **2012**, *13*, 770–780.
[6] R. Rodriguez, K. M. Miller, J. V. Forment, C. R. Bradshaw, M. Nikan, S. Britton, T. Oelschlaegel, B. Xhemalce, S. Balasubramanian, S. P. Jackson, *Nat. Chem. Biol.* **2012**, *8*, 301–310.
[7] S. M. Hampel, A. Pepe, K. M. Greulich-Bode, S. V. Malhotra, A. P. Reszka, S. Veith, P. Boukamp, S. Neidle, *Mol. Pharmacol.* **2013**, *83*, 470–480.
[8] R. Hänsel-Hertsch, D. Beraldi, S. V. Lensing, G. Marsico, K. Zyner, A. Parry, M. Di Antonio, J. Pike, H. Kimura, M. Narita, *Nat. Genet.* **2016**, *48*, 1267–1272.
[9] J. L. Mergny, J. F. Riou, P. Mailliet, M. P. Teulade-Fichou, E. Gilson, *Nucleic Acids Res.* **2002**, *30*, 839–865.
[10] Z. A. E. Waller, L. A. Howell, C. J. MacDonald, M. A. O'Connell, M. Searcey, *Biochem. Biophys. Res. Commun.* **2014**, *447*, 128–132.
[11] R. Rodriguez, S. Müller, J. A. Yeoman, C. Trentesaux, J.-F. Riou, S. Balasubramanian, *J. Am. Chem. Soc.* **2008**, *130*, 15758–15759.
[12] M. Di Antonio, G. Biffi, A. Mariani, E.-A. Raiber, R. Rodriguez, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2012**, *51*, 11073–11078; *Angew. Chem.* **2012**, *124*, 11235–11240.
[13] K. V. Diveshkumar, S. Sakrikar, S. Harikrishna, V. Dhamodharan, P. I. Pradeepkumar, *ChemMedChem* **2014**, *9*, 2754–2765.
[14] J. Dash, P. S. Shirude, S. Balasubramanian, *Chem. Commun.* **2008**, 3055.
[15] A. Chauhan, S. Paladhi, M. Debnath, J. Dash, *Org. Biomol. Chem.* **2016**, *14*, 5761–5767.
[16] M. Di Antonio, R. Rodriguez, S. Balasubramanian, *Methods* **2012**, *57*, 84–92.
[17] D. Koirala, S. Dhakal, B. Ashbridge, Y. Sannohe, R. Rodriguez, H. Sugiyama, S. Balasubramanian, H. Mao, *Nat. Chem.* **2011**, *3*, 782–787.
[18] K. I. E. McLuckie, Z. A. E. Waller, D. A. Sanders, D. Alves, R. Rodriguez, J. Dash, G. J. McKenzie, A. R. Venkitaraman, S. Balasubramanian, *J. Am. Chem. Soc.* **2011**, *133*, 2658–2663.
[19] S. Müller, S. Kumari, R. Rodriguez, S. Balasubramanian, *Nat. Chem.* **2010**, *2*, 1095–1098.
[20] S. Müller, G. D. Pantoş, R. Rodriguez, S. Balasubramanian, *Chem. Commun.* **2009**, 80–82.
[21] Z. A. E. Waller, P. S. Shirude, R. Rodriguez, S. Balasubramanian, *Chem. Commun.* **2008**, 1467–1469.
[22] A. Bugaut, K. Jantos, J.-L. Wietor, R. Rodriguez, J. K. M. Sanders, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2008**, *47*, 2677–2680; *Angew. Chem.* **2008**, *120*, 2717–2720.
[23] M. Bejugam, S. Sewitz, P. S. Shirude, R. Rodriguez, R. Shahid, S. Balasubramanian, *J. Am. Chem. Soc.* **2007**, *129*, 12926–12927.
[24] K. Jantos, R. Rodriguez, S. Ladame, P. S. Shirude, S. Balasubramanian, *J. Am. Chem. Soc.* **2006**, *128*, 13662–13663.
[25] S.-T. D. Hsu, P. Varnai, A. Bugaut, A. P. Reszka, S. Neidle, S. Balasubramanian, *J. Am. Chem. Soc.* **2009**, *131*, 13399–13409.
[26] Z. A. E. Waller, S. A. Sewitz, S.-T. D. Hsu, S. Balasubramanian, *J. Am. Chem. Soc.* **2009**, *131*, 12628–12633.
[27] L. Hahn, N. J. Buurma, L. H. Gade, *Chem. Eur. J.* **2016**, *22*, 6314–6322.
[28] D. P. N. Gonçalves, R. Rodriguez, S. Balasubramanian, J. K. M. Sanders, *Chem. Commun.* **2006**, 4685–4687.
[29] R. Rodriguez, G. D. Pantoş, D. P. N. Gonçalves, J. K. M. Sanders, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2007**, *46*, 5405–5407; *Angew. Chem.* **2007**, *119*, 5501–5503.
[30] Z. A. E. Waller, B. J. Pinchbeck, B. S. Buguth, T. G. Meadows, D. J. Richardson, A. J. Gates, *Chem. Commun.* **2016**, *52*, 13511–13514.
[31] A. Shivalingam, M. A. Izquierdo, A. L. Marois, A. Vyšniauskas, K. Suhling, M. K. Kuimova, R. Vilar, *Nat. Commun.* **2015**, *6*, 8178.
[32] S. A. Ohnmacht, S. Neidle, *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2602–2612.
[33] E. M. Rezler, J. Seenisamy, S. Bashyam, M.-Y. Kim, E. White, W. D. Wilson, L. H. Hurley, *J. Am. Chem. Soc.* **2005**, *127*, 9439–9447.
[34] T.-M. Ou, J. Lin, Y.-J. Lu, J.-Q. Hou, J.-H. Tan, S.-H. Chen, Z. Li, Y.-P. Li, D. Li, L.-Q. Gu, *J. Med. Chem.* **2011**, *54*, 5671–5679.
[35] S. Alcaro, A. Artese, G. Costa, S. Distinto, F. Ortuso, L. Parrotta, *Biochimie* **2011**, *93*, 1267–1274.
[36] S. A. Ohnmacht, C. Marchetti, M. Gunaratnam, R. J. Besser, S. M. Haider, G. Di Vita, H. L. Lowe, M. Mellinas-Gomez, S. Diocou, M. Robson, *Sci. Rep.* **2015**, *5*, 11385.
[37] G. Prato, S. Silvent, S. Saka, M. Lamberto, D. Kosenkov, *J. Phys. Chem. B* **2015**, *119*, 3335–3347.
[38] S. A. Ohnmacht, S. Neidle, *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2602–2612 ■■duplicate Ref. 32, please remove and adjust■■.
[39] S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug Discovery* **2011**, *10*, 261–275.
[40] S. Müller, R. Rodriguez, *Expert Rev. Clin. Pharmacol.* **2014**, *7*, 663–679.
[41] S. M. Hampel, A. Sidibe, M. Gunaratnam, J.-F. Riou, S. Neidle, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6459–6463.
[42] S. Mpima, S. A. Ohnmacht, M. Barletta, J. Husby, L. C. Pett, M. Gunaratnam, S. T. Hilton, S. Neidle, *Bioorg. Med. Chem.* **2013**, *21*, 6162–6170.
[43] G. W. Collie, R. Promontorio, S. M. Hampel, M. Micco, S. Neidle, G. N. Parkinson, *J. Am. Chem. Soc.* **2012**, *134*, 2723–2731■■duplicate Ref. 4, please remove and adjust■■.
[44] M. Islam, S. Fujii, S. Sato, T. Okauchi, S. Takenaka, *Molecules* **2015**, *20*, 10963–10979.

[45] The nitro-protected arginine was used for synthetic reasons: the unprotected derivative is insoluble in DMF and does not allow the synthesis of the corresponding NDI. Numerous attempts at the deprotection of NDI **5** failed, presumably owing to solubility problems.
[46] P. Pengo, G. D. Pantos, S. Otto, J. K. M. Sanders, *J. Org. Chem.* **2006**, *71*, 7063 – 7066.
[47] K. Tambara, N. Ponnuswamy, G. Hennrich, G. D. Pantoş, *J. Org. Chem.* **2011**, *76*, 3338 – 3347.
[48] N. Ponnuswamy, F. B. L. Cougnon, J. M. Clough, G. D. Pantoş, J. K. M. Sanders, *Science* **2012**, *338*, 783 – 785.
[49] H. Shao, J. Seifert, N. C. Romano, M. Gao, J. J. Helmus, C. P. Jaroniec, D. A. Modarelli, J. R. Parquette, *Angew. Chem. Int. Ed.* **2010**, *49*, 7688 – 7691; *Angew. Chem.* **2010**, *122*, 7854 – 7857.
[50] N. J. Greenfield, *Nat. Protoc.* **2007**, *1*, 2876 – 2890.
[51] A. I. S. Holm, L. M. Nielsen, S. V. Hoffmann, S. B. Nielsen, *Phys. Chem. Chem. Phys.* **2010**, *12*, 9581.
[52] N. Smargiasso, F. Rosu, W. Hsia, P. Colson, E. S. Baker, M. T. Bowers, E. De Pauw, V. Gabelica, *J. Am. Chem. Soc.* **2008**, *130*, 10208 – 10216.
[53] D. L. Ang, B. W. J. Harper, L. Cubo, O. Mendoza, R. Vilar, J. Aldrich-Wright, *Chem. Eur. J.* **2016**, *22*, 2317 – 2325.
[54] J. B. Chaires, D. Graves, Eds. , *Quadruplex Nucleic Acids*, Springer Berlin, Berlin, **2013**.
[55] A. I. S. Holm, B. Kohler, S. V. Hoffmann, S. Brøndsted Nielsen, *Biopolymers* **2010**, *93*, 429 – 433.
[56] K. R. Fox, Ed. , *Drug-DNA Interaction Protocols*, Humana Press, Totowa, NJ **2010**.
[57] H. Shao, M. Gao, S. H. Kim, C. P. Jaroniec, J. R. Parquette, *Chem. Eur. J.* **2011**, *17*, 12882 – 12885.
[58] M. Gao, S. Paul, C. D. Schwieters, Z.-Q. You, H. Shao, J. M. Herbert, J. R. Parquette, C. P. Jaroniec, *J. Phys. Chem. C* **2015**, *119*, 13948 – 13956.
[59] S. H. Kim, Y. Sun, J. A. Kaplan, M. W. Grinstaff, J. R. Parquette, *New J. Chem.* **2015**, *39*, 3225 – 3228.
[60] H. Shao, T. Nguyen, N. C. Romano, D. A. Modarelli, J. R. Parquette, *J. Am. Chem. Soc.* **2009**, *131*, 16374 – 16376.
[61] S. Tu, S. H. Kim, J. Joseph, D. A. Modarelli, J. R. Parquette, *J. Am. Chem. Soc.* **2011**, *133*, 19125 – 19130.
[62] S. H. Kim, J. A. Kaplan, Y. Sun, A. Shieh, H.-L. Sun, C. M. Croce, M. W. Grinstaff, J. R. Parquette, *Chem. Eur. J.* **2015**, *21*, 101 – 105.

Manuscript received: March 1, 2017
Accepted Article published: April 13, 2017
Final Article published: ■■ ■■, 0000

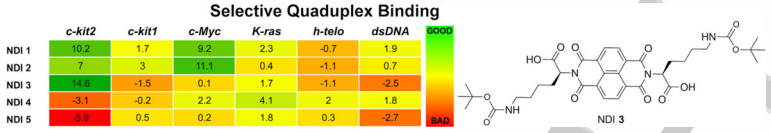
FULL PAPER

DNA recognition

*D. M. Răsădean, B. Sheng, J. Dash, G. D. Pantoş**

■■ – ■■

Amino-Acid-Derived Naphthalenediimides as Versatile G-Quadruplex Binders



Selective oncogene stablilisation: Water-soluble, amino acid functionalised naphthalenediimides are potential ligands of native G-quadruplexes, and they are tested on a panel of oncogene promoters and against double-strand DNA. One of the ligands, NDI **3**, exhibits a highly discriminating nature by stabilising only the oncogene promoter *c-kit2*, which is overexpressed up to 80% in ovarian, gastrointestinal, and breast malignancies (see figure).

Pantos et al (@UniofBath) evaluated #naphthalenediimides as ligands for #Gquadruplexes and #oncogene promoters SPACE RESERVED FOR IMAGE AND LINK

Share your work on social media! *Chemistry - A European Journal* has added Twitter as a means to promote your article. Twitter is an online microblogging service that enables its users to send and read text-based messages of up to 140 characters, known as "tweets". Please check the pre-written tweet in the galley proofs for accuracy. Should you or your institute have a Twitter account, please let us know the appropriate username (i.e., @accountname), and we will do our best to include this information in the tweet. This tweet will be posted to the journal's Twitter account @ChemEurJ (follow us!) upon online publication of your article, and we recommended you to repost ("retweet") it to alert other researchers about your publication.

Please check that the ORCID identifiers listed below are correct. We encourage all authors to provide an ORCID identifier for each coauthor. ORCID is a registry that provides researchers with a unique digital identifier. Some funding agencies recommend or even require the inclusion of ORCID IDs in all published articles, and authors should consult their funding agency guidelines for details. Registration is easy and free; for further information, see http://orcid.org/.

Dora M. Răsădean
Bin Sheng
Dr. Jyotirmaee Dash
Dr. G. Dan Pantoş http://orcid.org/0000-0003-2200-550X